

## Sentiment Analysis & Natural Language: Processing Techniques for Capital Markets & Disclosure

By *Nicolas H.R. Dumont*

Application of machine learning, artificial intelligence, and advanced analytics to big data influences nearly all industries today. The securities markets are no exception. Issuers use these tools for marketing, product development, and operations; investors harness data in search of trading insights; and regulators monitor compliance and detect risks. Markets themselves create a wealth of data that perpetuates a virtuous cycle of information generation, reliance and analysis.

To date, discussions surrounding technology and finance focus on how each player uses innovation to achieve its own goals, whether streamlining operations, increasing returns, or detecting fraud.<sup>1</sup> There is considerably less discussion about how these developments born of the Internet era influence (or should influence) issuers. Following an overview of sentiment analysis technology and how and to what extent it is currently being used in the capital markets, we then discuss the way these techniques could affect how issuers operate in the market.

### The State of Play

#### “Big Data”

Business and finance have both long relied on data, as the term is used in the casual sense. A standard definition of data is any “factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation.”<sup>2</sup> As implied by the terminology, big data is partially characterized by volume. But beyond sheer size, “big data” connotes a degree of complexity that results from

the compilation of information taking different forms, stemming from multiple sources, produced at different times.<sup>3</sup> This complexity and volume calls for techniques that enable the capture, storage, processing, and analysis of such information. In this way, “big data” is more than just *a lot* of information; it represents a new frame in which information is collected, connected, and used.

While investors have long sought out new forms of information to enhance returns, the digital age generates exponentially more data from countless new sources. Investors can harness satellite images to measure customer cars in parking lots, and can “scrape” issuer Web sites for more information than what was perhaps intended for public consumption. In fact, data is now so big—in terms of relevance and scope—that many investors, and even some government agencies,<sup>4</sup> subscribe to data feeds to be processed in house, or prepackaged analytics from data analytics companies.<sup>5</sup>

### Natural Language Processing & Sentiment Analysis

The proliferation of big data has required and encouraged new processing methods, and new methods have in turn required and encouraged new data sources. As the sheer amount of information grows and becomes more complex, storage and processing techniques become increasingly important, but as the universe of data constantly grows and evolves, it is increasingly inefficient and ineffective to rely on predetermined programming to govern processing techniques. A new area of artificial intelligence, known broadly as machine learning, responds to this issue. Such algorithms not only analyze data but also use such data to learn and enhance processing rules such that they adapt and change without additional guidance.<sup>6</sup>

© 2017 Gibson, Dunn & Crutcher LLP. Nicolas H.R. Dumont is an Associate in the New York office of Gibson, Dunn & Crutcher LLP.

---

Natural language processing (NLP) developed in response to yet a third issue presented by big data. Much of the information that is traditionally important in capital markets is unstructured, meaning it is formatted and designed for humans, not computers, such as Management's Discussion and Analysis (MD&A) disclosures, financial footnotes, and oral disclosures. Applying machine-learning techniques to spoken and written language, NLP algorithms process these portions, and other sources, and learn to read and interpret language.<sup>7</sup>

For example, algorithms can now automatically supplement structured financial disclosures from Securities and Exchange Commission (SEC) filings with information from the textual disclosures without an analyst actually reading the text and manually adjusting a model.<sup>8</sup> And as NLP has developed, algorithms have advanced from mere text retrieval to automatic categorization and topic modelling, such that NLP algorithms can now retrieve filings, financial reports, press releases, news and the like, and then also compare the various sources to verify consistency, detect differences, synthesize information, and incorporate a wider range of sources into analysts' models.<sup>9</sup>

To illustrate the speed at which NLP can operate, consider Twitter's experience in April 2015 when that company accidentally posted an earnings report an hour early: A Web crawler using NLP algorithms seized the report, summarized, and then tweeted the contents within three seconds of the post.<sup>10</sup> In the regulatory space, NLP enables the SEC to rapidly scan regulatory filings and discover new terms as they appear, potentially unveiling new risks to the market as a whole or to specific industries.<sup>11</sup> Other NLP applications use sentence length or complexity as a proxy for obfuscation, which in turn measures risk,<sup>12</sup> and compares the detail or length of the topics discussed in MD&As as a proxy for accounting fraud.<sup>13</sup>

NLP also encourages the use of new data sources, now accessible through improved techniques. Algorithms mine less traditional sources

of information, such as news and social media feeds, for insights on consumer trends and to discover market-moving events. For example, Dataminr, a company that sells real-time alerts based on social media feeds, alerted clients when the King of Saudi Arabia died more than four hours before crude oil prices spiked.<sup>14</sup> Similarly, using thousands of user posts on a Reddit feed, Eagle Alpha—a software company that provides data and analysis tools—predicted that Electronic Arts would sell more copies of its new video game than originally predicted before the company revised its projections.<sup>15</sup> Thus, social media and other alternative sources have earned the respect of at least some investors who seek to harness both the wisdom of the crowds and the speed at which news travels in these networks.

Finally, investors and the SEC also use NLP in sentiment analysis, a tool to assess issuer or consumer outlook. Premised on the idea that particular words connote uncertainty, intentional obfuscation, or a positive or negative outlook, investors and regulators alike use algorithms to measure prevalence of certain words,<sup>16</sup> and then draw inferences based on these subtle indicators. Positive or negative sentiment scores not only synthesize sizable amounts of language into a single composite score,<sup>17</sup> but can also be applied to portions of texts to show that optimism in a portion of a disclosure is camouflaging uncertainty in another.<sup>18</sup> The MD&A portion of a quarterly or annual report is particularly prone to sentiment analysis, as such disclosures are required for public issuers, the topics are dictated, and the disclosure is explicitly geared toward measuring management's perspective.<sup>19</sup> However, sentiment analysis is also applied to derive tonality from other corporate sources, such as oral statements on earnings calls<sup>20</sup> and press releases, as well as consumer sources, like user reviews, social media, and news.<sup>21</sup>

## Applications & Challenges

Both investors and regulators are increasingly applying these new techniques to achieve their respective goals of higher returns and enforcement of market rules and regulations.<sup>22</sup>

---

## *Use by Traders*

One estimate shows that more than a quarter of stock turnover is traded by funds run by algorithms using these techniques, up 100 percent in 2017 over just four years prior.<sup>23</sup> Half of the top 25 investment firms rely on a computer-based strategy of some kind.<sup>24</sup> BlackRock's AI machine, Aladdin, uses NLP to sift through sources from broker reports to social media feeds to generate sentiment scores and learn about news events, and Bridgewater Associates uses IBM Watson technology to glean insights on and predict market trends.<sup>25</sup>

Machine learning is still learning, however, and flaws remain. First, even as sentiment analysis improves, sarcasm or other non-“plain English” text can pose significant interpretation challenges for machines.<sup>26</sup> Further, while machines easily detect correlations, it is significantly harder to learn causality, which limits the application of the derived insights.<sup>27</sup> Especially given the large amount of data incorporated, models are bound to produce at least some strong correlations based on historical trends that are not representative of actual relationships that will hold in the future. Some correlations are just coincidental.

Lastly, increased reliance on algorithms also leaves investors vulnerable to false positives. For example, algorithms—and possibly people, as well—were fooled by two fake tweets sent under handles designed to pass for well-known market players, and a hacker manipulating the Associated Press's Web site sent the market down 145 points in only two minutes.<sup>28</sup>

For some analysts, the use of algorithms to interpret market data appears to have dampened the traditional market-moving effects of released material information. Some have attributed a recent decrease in market volatility to the rise of advanced analytical methods that integrate many small indicators rather than react to “material” information distributed by issuers. Over half of the lowest 25 volatility readings, as based on an options based index called VIX, were observed between May and July 2017.<sup>29</sup> Analytics experts attribute this lower volatility

to algorithms that process news and events over a longer period, spreading the impact of what could be material information over a longer trading period.<sup>30</sup> An alternative theory is that the algorithms access and process so much data that models are converging, reducing spreads and the associated volatility.

There are reasons to be skeptical of this view, however. Data proliferation when combined with automated trading software creates risks, especially when that information is replaced by unverified outside sources. For example, human subscribers to the Muddy Waters and Citron Research feeds would not have been fooled by the fake tweets sent out in 2013 because they would compare the information to the real, vetted source.<sup>31</sup> Similarly, combining these less-vetted sources with processing systems that few understand can also downplay truly material information and focus too much attention on the noise. Synthesizing machine-simplified disclosures with indicators collected from disparate sources risks replacing deliberate nuances with random, potentially misleading ones. Further, in a world in which information is not only reported, tweeted, and posted, but also then re-tweeted and re-posted, algorithms risk mistaking echo chambers for trends. These issues, which have occasionally manifested themselves in actual volatility, call for increased monitoring by issuers of information relating to them.

## *Use by the SEC*

The SEC, like many investors,<sup>32</sup> uses machine learning as a tool to assist, but not to replace, human judgment. In the enforcement context, the SEC uses transaction data and other information to detect insider trading, market manipulation, and compliance with suitability rules.<sup>33</sup> The SEC also employs NLP algorithms to sift through filings and discover new terms that could signal new risks and market exposures.<sup>34</sup> Both of these applications enable the SEC to take a more proactive approach to detecting fraud and risk. Rather than wait for suspicious behavior to be reported, the SEC can use modeling to uncover discrepancies. Even when the SEC does rely on tips, complaints, and referrals

---

submitted to the Office of the Whistleblower, NLP technology is used to group common complaints, enabling a more comprehensive and speedy review.<sup>35</sup>

Additionally, the SEC also uses sentiment analysis to assess the tonality of disclosures, looking for obfuscation or negativity as a measure of risk. Comparing the NLP results with more established risk indicators, like past enforcement actions or examination results, the SEC can more easily assess future filings by issuers known to pose risk, and train its models to aid examiners in deciding which other issuers deserve more scrutiny.<sup>36</sup> The SEC also compares the length and detail of disclosures for indicators of when obfuscation or brevity is a signal of fraud.

For example, the Division of Economic Risk and Analysis, using NLP techniques, has found that firms subject to enforcement actions related to financial reporting are less likely to discuss certain topics related to performance, essentially confirming that issuers charged with misconduct tend to downplay risks in financial disclosures.<sup>37</sup> In addition, regulators have started to address some of the implications for asset managers and funds, particularly as applied to computer-based strategy and robo-trading.<sup>38</sup>

## Implications & Recommendations

These developments represent a change in the way that the market digests financial information. Not only are investors and investigators alike regularly harnessing record amounts of data, but they are also increasingly looking at innovative ways to integrate such data into predictive methodologies. Consideration must be given to what these technological developments imply for issuers as a practical matter. A few thoughts are presented.

- *Everything is (likely) being monitored.* Earnings calls have long been market-moving events, but the advent of sentiment analysis means that investors might be listening in a new way. Issuers should be aware that written and oral statements are exposed to such

analysis and other NLP techniques, and that the unstructured or oral nature of a disclosure does not necessarily protect the content from machine analytics. As sentiment and topic variation are more easily detected in less structured settings, issuers should devote increased attention to the preparation and rehearsal of earnings calls, employee disclosures outside the firm, free writing prospectuses, 8-Ks and other less-scripted events. These disclosures should be reviewed carefully prior to dissemination with a view toward the way an algorithm designed to scrape information could interpret and retranscribe them, as nuance may be lost.

- *Consistency across and within disclosures.* Issuers should ensure absolute consistency between different forms of disclosure.
  - Technology enables market watchers to spot even the most minor discrepancies among statements based on length, detail, and clarity of content. While explicitly scripting a message to counteract a sentiment-analysis algorithm could be considered manipulative or misleading (and potentially a violation of Regulation FD if designed to signal information to sophisticated market participants), developing an understanding how these algorithms function is recommended.
  - Issuers should recognize that the SEC (and perhaps investors) compare the length of topical discussions within issuer filings or statements over time and across the filings of multiple issuers in search of evidence that issuers are attempting to camouflage or downplay risks.<sup>39</sup> While there are clearly instances in which an issuer can and should omit discussion of specific risks, it should be remembered that subtle differences will likely be noticed and may increase scrutiny. For instance, if an issuer consistently discusses an aspect of the business or a risk in a certain level of detail, regulators and the market will notice when the length or detail of such discussion changes. Similarly, if one issuer

---

omits a risk or condition discussed by its peers, market participants and watch dogs will also take notice. This has generally always been true, but NLP has enabled investors and regulators to notice more detail and notice it more quickly. Thus, investors should be prepared to explain such changes, even if minimal.

- *Reinforce Regulation FD-type controls.* While most public issuers have controls and policies in place to monitor what is said by or on behalf of the company, the new paradigm of data analysis leaves little room for error. As online communications, just like oral disclosures, are governed by Regulation FD, issuers should have clear policies in place to govern who can transmit information online, when the information can be posted, and how the company and its employees conduct Regulation FD analysis on an *ad hoc* basis. Issuers should also consider developing compliance procedures for tracking employee and company sites, as well as a record-retention policy. Issuers should also enforce the appropriate level of security for all publicly facing sites to ensure that material information is not inadvertently exposed. Twitter's experience in 2015 proves that accidental exposures or similar lapses in security are not overlooked by software scanning the market automatically.
  - Issuers should also be mindful to monitor (or altogether avoid) unscripted oral communications made by company officials, especially those made in private or informal settings, as it is possible that those utterances will be captured, scrutinized and analyzed quickly in the future. While in the past it would have been unusual for such statements to reach investor or regulatory scrutiny simply because proof was difficult to come by, an era in which voice recordings are transcribed onto permanent records may soon emerge.
- *Survey & control your digital footprint.* Because investors are scrutinizing far more data than that reported to the SEC, issuers should be attentive to that data over which they may

exercise control, and be otherwise aware of their pronouncements and their implications. Failure to monitor appropriately could result in exposure to increased liability.

- First, companies may consider reformatting portions of their Web site not intended to convey investor information in a manner that is less conducive to analysis by even the smartest machines. For example, retail issuers may consider how much inventory information is available through consumer shopping portals, as investors have developed Web crawlers that access retailer sites for information on prices and sales.<sup>40</sup> Some online platforms require verification before users can access certain information. Whether requiring a user to check a box or decode a message, such preliminary screening tools might deter some Web crawlers from extracting data that is not intended to be used for trading.
- Second, especially in the absence of any Regulation FD changes, issuers should consider their own data-sharing arrangements. Many companies share or sell data about the company or its customers, which in turn, implies information about the company. For instance, payment processors, like banks and credit cards, often negotiate the rights regarding information gleaned from consumer transaction data. The terms of these arrangements should be considered, not only with respect to consumer privacy and marketing potential, but also in light of the new uses to which data has been put. For example, for an issuer who sells consumer transaction data aggregated by issuer rather than customer demographic, the data could easily now convey sensitive financial information about the issuer, rather than its customers, as intended.
  - Issuers that do share data should also conduct an appropriate Regulation FD analysis tailored to their own facts and circumstances to determine whether the receiving parties are covered under Regulation FD, whether the

---

information is material, and whether the data is released on a schedule that could conflict with disclosure rules.

- In sum, given the prevalence of data in markets, companies should review data contracts to see how much information they are sharing, when it is transmitted, what that information can be used for, and with whom that data can be shared.
- Third, because information and new stories can be generated quickly and without filter or scrutiny, issuers should continue to monitor the proliferation of stories for which they are not responsible that are nevertheless erroneous, and act swiftly to correct the story (if possible) before it is reproduced too quickly or causes any movements in the markets.
- Fourth, while the SEC encourages issuers to use company Web sites as a method of communication, issuers should also recognize that these less formal environments are likely to attract equal if not more attention from investors. Just as the MD&A portion of a filing is a ripe target for sentiment analysis because it is less structured, Web sites and social media often offer additional opportunities to mine executive and company statements for subtle clues regarding outlook and future performance. Issuers should ensure that all statements are carefully drafted, even when conversing orally or posting in a seemingly more casual environment.
- *Consider what constitutes “material” information.* What constitutes “material” information in the age of machine learning? When the US Securities Act was enacted in 1933, information regarding issuers was difficult to access and far harder to distribute than today. The disclosure regime imposed by U.S. securities laws effectively contemplates an information pipeline: Issuers determine what information could be considered material to investors in response to forms developed, and events identified, by

the US SEC. While it has always been true that investors have sought to gain an advantage in the markets (sometimes illegally) by looking outside the “pipeline,” new data, whether oral, written, or on social media, exacerbates the chatter.

- Advances in big data and analytics call into question an approach to material disclosure based solely on the opinions of human drafters. Because company disclosures and other information outside the “pipeline” are analyzed to decipher and discover hidden meanings not contemplated by, or hidden from, their writers, “material” information may be more difficult to identify from the perspective of an issuer. The securities laws were designed to create information digestible by the typical investor, who would presumably review and analyze such information in a similar manner one to the next.

Today, in reality, information is now digested, interpreted, and acted upon by algorithms at speeds exceeding human capacity. Those algorithms are, in certain circumstances, making investment decisions that almost by definition make unseen data (and related patterns in that data) material. It is unlikely that these interpretations, at least in all instances, are what the typical issuer intends when accounting, investor relations, and legal teams produce disclosures for the investing public. With limited data and limited tools, crafting nuanced disclosures to be read by humans produces a certain type of disclosure that has long guided the public markets. With lots of data that is easily produced and distributed, relating what is “material” may be more challenging because counterparties are listening in a way that most humans never intended.

## Notes

1. See, e.g. Penny Crosman, “All the Ways AI Will Slash Wall Street Jobs,” *Am. Banker*, Mar. 16, 2017, available at <https://www.americanbanker.com/news/all-the-ways-ai-will-slash-wall-street-jobs> (noting Wall Street’s use of AI not

---

only in trading but also in robotic process automation, front office work, compliance and HR), last accessed Oct. 10, 2017; Bryan Yurcan, "What Santander's Latest Bets Say About the Future of Fintech," *Am. Banker*, Jul. 12, 2017, available at <https://www.americanbanker.com/news/what-santanders-latest-bets-say-about-the-future-of-fintech> (highlighting use of real time sentiment analysis of phone calls to aid call center workers) last accessed Oct. 10, 2017.

2. *Data*, Merriam-Webster Dictionary (online ed.), <https://www.merriam-webster.com/dictionary/data>, last accessed Oct. 11, 2017.

3. See Svetlana Sicular, Contributor, "Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three 'V's,'" *Forbes Mag.*, Mar. 27, 2013, available at <http://onforb.es/103sM27>, last accessed Oct. 10, 2017.

4. Jen Wieczner, "How Investors are Using Social Media to Make Money," *Fortune Mag.* (Dec. 7, 2015), available at <http://fortune.com/2015/12/07/dataminr-hedge-funds-twitter-datal>, last accessed Oct. 10, 2017.

5. See Sarah Butcher, "43 of the Top Big Data Companies to Work For, by J.P. Morgan," *efinancialcareers* (Jun. 1, 2017), available at <http://news.efinancialcareers.com/us-en/285462/the-top-big-data-companies-to-work-for-by-j-p-morgan/> (listing top financial data services) last accessed Oct. 10, 2017; Ivy Schmerken, "Quant Funds Get Sentimental About Big Data," *Market Media*, Aug. 4, 2016, available at <https://marketsmedia.com/quant-funds-sentimental-big-data-by-ivy-schmerken-flextrade/> (discussing various analysis services), last accessed Oct. 10, 2017; Penny Crosman, "Beyond Robo-Advisers: How AI Could Rewire Wealth Management," *Am. Banker*, Jan. 5, 2017, available at <https://www.americanbanker.com/news/beyond- robo-advisers-how-ai-could-rewire-wealth-management> (discussing in house analytics), last accessed Oct. 10, 2017.

6. Tom Loftus, "What Your CEO Is Reading: Designing Around AI's 'Black Box'; Flash Organizations; Nature's End," *The Wall St. J.*, Jul. 14, 2017, available at <http://on.wsj.com/2tTvvV3> (quoting Katharine Schwab from Fast Company), last accessed Oct. 10, 2017 (subscription required).

7. Alex LaPlante and Thomas F. Coleman, *Teaching Computers to Understand Human Language: How Natural Language Processing is Reshaping the World of Finance*, The Global Risk Institute, Jan. 15, 2017, available at <http://globalriskinstitute.org/publications/natural-language-processing-reshaping-world-finance/>, last accessed Oct. 10, 2017.

8. "Achieving Transparency in Financial Reporting With Artificial Intelligence," *Kognetics Blog* (Oct. 3, 2016), available at <http://www.kognetics.com/2016/10/3/achieving-transparency-in-financial-reporting-with-artificial-intelligence/> last accessed Oct. 10, 2017.

9. LaPlante and Coleman, *supra* n.7.

10. Wieczner, *supra* n.4.

11. Scott W. Bauguess, Champagne Keynote Address: The Role of Big Data, Machine Learning, and AI in Assessing Risks: a Regulatory Perspective (June 21, 2017), available at <https://www.sec.gov/news/speech/bauguess-big-data-ai>, last accessed Oct. 10, 2017.

12. See David F. Larcker and Anastasia A Zakolyukina, "Detecting Deceptive Discussions in Conference Calls," 50 *J. of Acct. Res.* 495, 499 (2012) (citing other sources).

13. Gerard Hoberg and Craig Lewis, "Do Fraudulent Firms Produce Abnormal Disclosure?," 43 *J. of Corp. Fin.* 58, 75-77 (2017) (showing which topics are correlated with a fraud finding when measured by deviation from a standard explanation).

14. Wieczner, *supra* n.4.

15. *Id.*

16. A great deal of academic research focuses on the proper mapping of words to sentiment. See generally Sanjib Ranjan Das, "Text and Context: Language Analytics in Finance," 8 *Found. and Trends in Fin.* 3 (2014); Ann Devitt and Khurshid Ahmad, "Is There a Language of Sentiment? An Analysis of Lexical Resources For Sentiment Analysis," 47 *Language Resources and Evaluation* 475 (2013); Jessen L Hobson, William J. Mayew and Mohan Venkatachalam, "Analyzing Speech to Detect Financial Misreporting," 50 *J. of Acct. Res.* 349 (2012); Tim Loughran and Bill McDonald, "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," 66 *J. of Fin.* 35, (2011).

17. Andreas Chouliaras, "The Pessimism Factor: SEC EDGAR Form 10-K Textual Analysis and Stock Returns" (Jul. 6, 2017), available at <http://dx.doi.org/10.2139/ssrn.2627037>, last accessed Oct. 10, 2017; LaPlante and Coleman, *supra* n.7; Schmerken, *supra* n.5.

18. Hoberg and Lewis, *supra* n.13 at 76-79.

19. *Id.* at 59.

20. Larcker and Zakolyukina, *supra* n.12 at 499.

21. Schmerken, *supra* n.5; Nigel Farmer, "Sentiment Analysis: Where Next?," *Markets Media*, (Sept. 11, 2015), available at <https://marketsmedia.com/sentiment-analysis-where-next/>, last accessed Oct. 10, 2017.

22. The SEC has been increasing its use of machine learning in enforcement and compliance as well as risk detection, as discussed later. Additionally, it is reported that Dataminr, a data and analysis supplier to large investment firms, also has government clients. See Wieczner, *supra* n.4.

23. Spencer Jakab, "How Quants Are Calming the Stock Market," *The Wall St. J.*, B12, July 21, 2017 (citing Tabb Group).

24. Jonathan Ratner, "Last days of the stock picker as money managers embrace artificial intelligence," *Fin. Post*, Apr. 7, 2017, available at <http://business.financialpost.com/>

---

*investing/last-days-of-the-stock-picker-as-money-managers-embrace-artificial-intelligence/wcm1333a7ba8-e0c9-487f-8787-2cb67970300e#comments-area*, last accessed Oct. 10, 2017.

25. Crosman, *supra* n.5.

26. Penny Crosman, “Why AI Still Has a Ways to Go in Wealth Management,” *Am. Banker*, Jul. 13, 2017, available at <https://www.americanbanker.com/news/why-ai-still-has-a-ways-to-go-in-wealth-management>, last accessed Oct. 10, 2017; Wieczner, *supra* n.4 (noting that one hedge fund’s AI was fooled by sarcastic tweets regarding Lululemon).

27. Ratner, *supra* n.24.

28. Wieczner, *supra* n.4.

29. Jakab, *supra* n.23.

30. *Id.*

31. Wieczner, *supra* n.4 (discussing how an individual set up imposter accounts to tweet fake news about companies Audience and Sarepta Therapeutics, causing the stocks to fall 28 percent and 16 percent respectively).

32. Crosman, *supra* n.26.

33. See Mary Jo White, Chair, A New Model for SEC Enforcement: Producing Bold and Unrelenting Results

(Nov. 18, 2016), available at <https://www.sec.gov/news/speech/chair-white-speech-new-york-university-111816.html>, last accessed Oct. 10, 2017.

34. Bauguess, *supra* n.11.

35. *Id.*

36. Scott W. Bauguess, Has Big Data Made Us Lazy? (Oct. 21, 2016), available at <https://www.sec.gov/news/speech/bauguess-american-accounting-association-102116.html>, last accessed Oct. 10, 2017.

37. Bauguess, *supra* n.11.

38. Financial Stability Board, *Financial Stability Implications from FinTech: Supervisory and Regulatory Issues that Merit Authorities’ Attention* (Jun. 27, 2017) (focusing on robo-advisors and machine learning); International Organization of Securities Commissions, *IOSCO Research Report on Financial Technologies* (Feb. 2017) (discussing risks from robo-advisors and social media analysis, among other topics); Financial Industry Regulatory Authority, *Report on Digital Investment Advice* (March 2016) (analyzing digital investment advice through traditional broker-dealer rules).

39. Bauguess, *supra* n.11.

40. Wieczner, *supra* n.4.